Exhibit 25

Chapter 14

From Association to Causation: Deriving Inferences from Epidemiologic Studies

Not everything that can be counted counts, and not everything that counts can be counted.

—William Bruce Cameron, 1963[1]

Learning Objectives

- To describe a frequent sequence of study designs used to address questions of etiology in human populations.
- To differentiate between real and spurious associations in observational studies.
- To define necessary and sufficient in the context of causal relationships.
- To present guidelines for judging whether an association is causal based on the guidelines set forth by the U.S. Surgeon General, and to discuss the application of these guidelines to broader questions of causal inference.
- To describe how the guidelines for causation originally proposed by the Surgeon General have been modified and utilized by the U.S. Public Health Service and the U.S. Preventive Services Task Force.

In the previous chapters, we discussed the designs of epidemiologic studies that are used to determine whether an association exists between an exposure and a disease (Fig. 14-1A). We then addressed different types of risk measurement that are used to quantitatively express an excess in risk. If we determine that an exposure is associated with a disease, the next question is whether the observed association re! ects a causal relationship (Fig. 14-1B).

Although Figures 14-1A and B refer to an environmental exposure, they could just as well have speci! ed a genetic characteristic or characteristics or a speci! c combination of environmental and genetic factors. As we shall see in Chapter 16, studies of disease etiology generally address the contributions of both genetic and environmental factors and their interactions.

This chapter discusses the derivation of causal inferences in epidemiology. Let us begin by asking, "What approaches are available for studying the etiology of disease?"

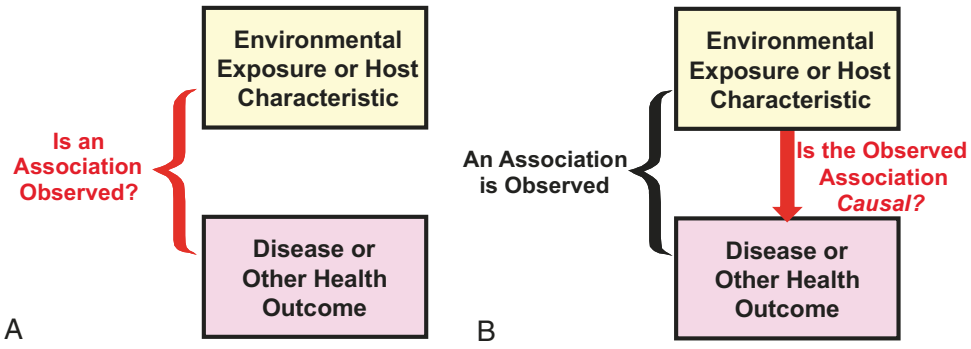


Figure 14-1. **A,** Do we observe an association between exposure and disease? **B,** Is the observed association between exposure and disease causal?

APPROACHES FOR STUDYING DISEASE ETIOLOGY

If we are interested in whether a certain substance is carcinogenic in human beings, a ! rst step in the study of the substance's effect might be to expose animals to the carcinogen in a controlled laboratory environment. Although such animal studies afford us the opportunity to control the exposure dose and other environmental conditions and genetic factors precisely, and to keep loss to follow-up to a minimum, at the conclusion of the study we are left with the problem of having to extrapolate data across species, from animal to human populations. Certain diseases seen in humans have neither occurred nor been produced in animals. It is also dif! cult to extrapolate animal doses to human doses, and species differ in their responses. Thus, although such toxicologic studies can be very useful, they still leave a gnawing uncertainty as to whether the animal ! ndings can be generalized to human beings.

We can also use in vitro systems, such as cell culture or organ culture. However, because these are arti! cial systems, we are again left with the dif- ! culty of extrapolating from arti! cial systems to intact, whole human organisms.

In view of these limitations, if we want to be able to draw a conclusion as to whether a substance causes disease in human beings, we need to make *observations in human populations*. Because we cannot ethically or practically randomize human beings to exposure to a suspected carcinogen, we are dependent on nonrandomized observations, such as those made in case-control and cohort studies.

Approaches to Etiology in Human Populations

Epidemiology capitalizes on what have been called "unplanned" or "natural" experiments. (Some think that this phrase is a contradiction in terms, in that the word "experiment" implies a planned exposure.) What we mean by *unplanned* or *natural* experiments is that we take advantage of groups of people who have been exposed for nonstudy purposes, such as occupational cohorts in speci! c industries or persons exposed to toxic chemicals. Examples include people affected by the poison gas leak disaster at a pesticide manufacturing plant in Bhopal, India, in 1984 and residents of Hiroshima and Nagasaki, Japan, who were exposed to radiation from the atomic bombs dropped on both cities in 1945. Each of these exposed groups can be compared to a nonexposed group to determine whether there is an increased risk of a certain adverse effect in persons who have been exposed.

In conducting human studies, the sequence shown in Figure 14-2 is frequently followed:

The initial step may consist of *clinical observations* at the bedside. For example, when the surgeon Alton Ochsner observed that virtually every patient on whom he operated for lung cancer gave a history of cigarette smoking, he was among the ! rst to suggest a possible causal relationship.[2] A second step is to try to identify *routinely available data*, the analysis of which might shed light on the question. We can then carry out *new studies* such as the cohort and case-control studies discussed in Chapters 9 and 10, which are speci! cally designed to determine whether there is an association between an exposure and a disease, and whether a causal relationship exists.

The usual ! rst step in carrying out new studies to explore a relationship is often a *case-control study*. For example, if Ochsner had wanted to further explore his suggestion that cigarette smoking may be associated with lung cancer, he would have compared the smoking histories of a group of his patients with lung cancer with those of a group of patients without lung cancer—a case-control study.

If a case-control study yields evidence that a certain exposure is suspect, we might next do a *cohort study* (e.g., comparing smokers and nonsmokers and determining the rate of lung cancer in each group or comparing workers exposed to an

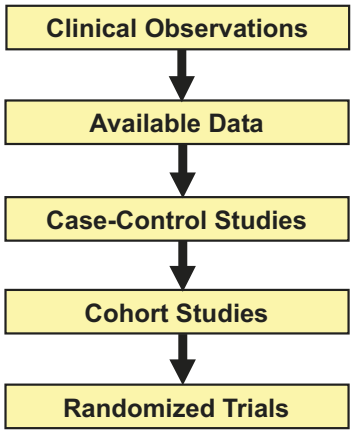


Figure 14-2. A frequent sequence of studies in human populations.

Figure 14-3. Another example of association or causation. (DILBERT © 2011 Scott Adams. Used by permission of UNIVERSAL UCLICK. All rights reserved.)

industrial toxin with workers without such an exposure). Although, in theory, a randomized trial might be the next step, as discussed earlier, randomized trials are almost never used to study the effects of putative toxins or carcinogens and are generally used only for studying potentially beneficial agents.

Conceptually, a two-step process is followed in carrying out studies and evaluating evidence. However, in practice, this process often becomes interactive and deviates from a fixed sequence:

1. We determine whether there is an association or correlation between an exposure or characteristic and the risk of a disease (Fig. 14-3). To do so, we use:
   a. Studies of group characteristics: ecologic studies (discussed in Chapter 10, p. 208)
   b. Studies of individual characteristics: cohort, case-control, and other types of studies
2. If an association is demonstrated, we determine whether the observed association is likely to be a causal one.

TYPES OF ASSOCIATIONS

Real or Spurious Associations

Let us turn next to the types of associations that we might observe in a cohort or case-control study. If we observe an association, we start by asking the question, "Is it a true (real) association or a false (spurious) one?" For example, if we designed a study to select controls in such a way that they tended to be nonexposed, we might observe an association of exposure with disease (i.e., more exposure in cases than in controls). This would not be a true association, but only a result of the study design. Recall that this issue was raised in Chapter 10 regarding a study of coffee consumption and cancer of the pancreas. The possibility was suggested that the controls selected for the study had a lower rate of coffee consumption than was found in the general population.

Interpreting Real Associations

If the observed association is real, is it causal? Figure 14-4 shows two possibilities. Figure 14-4A shows a causal association: we observe an association of exposure and disease, as indicated by the bracket, and the exposure induces development of the disease, as indicated by the arrow. Figure 14-4B shows the same observed association of exposure and disease, but they are associated only because they are both linked to a third factor, designated here as *factor X*. This association is a result of confounding and is noncausal. Confounding is discussed in greater detail in Chapter 15.

In Chapter 10 we discussed this issue in relation to McMahon's study of coffee and cancer of the

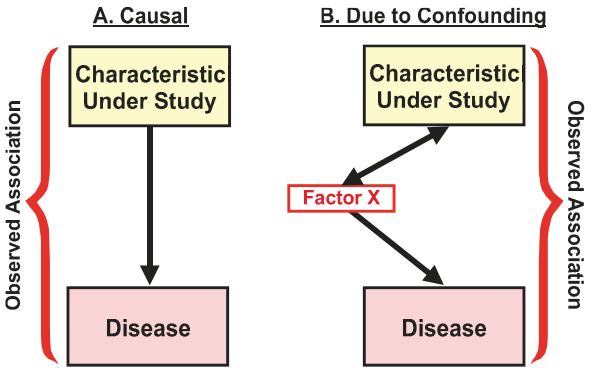


Figure 14-4. Types of associations.

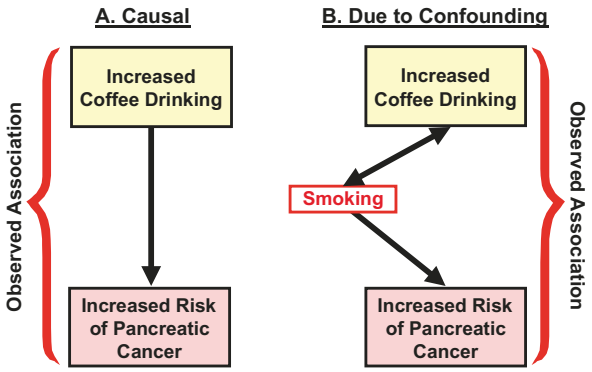


Figure 14-5. Interpreting an observed association between increased coffee drinking and increased risk of pancreatic cancer.

pancreas. McMahon observed an association of coffee consumption with risk of pancreatic cancer. Cigarette smoking was known to be associated with pancreatic cancer, and coffee drinking and cigarette smoking are closely associated (few smokers do not drink coffee) (Fig. 14-5). Therefore, was the observed association of coffee drinking and cancer of the pancreas likely to be a causal relationship, or could the association be due to the fact that coffee and cigarette smoking are associated, and that cigarette smoking is a known risk factor for cancer of the pancreas?

The same issue is exempli! ed by the observed association of increased serum cholesterol level and risk of coronary heart disease (CHD) (Fig. 14-6). Is increased cholesterol a causal factor for increased risk of CHD, or is the observed association due to confounding? That is, are we observing an association of increased cholesterol and CHD because both are associated with a factor X (such as a particular genetic pro! le), which might cause people to have both increased levels of cholesterol and an increased risk of CHD?

Is this distinction really important? What difference does it make? The answer is that it makes a tremendous difference from both clinical and public health standpoints. If the relationship is causal, we will succeed in reducing the risk of CHD if we lower cholesterol levels. However, if the relationship is due to confounding, then the increased risk of CHD is caused by factor X. Therefore, changes in the level of serum cholesterol will have no effect on the risk of CHD. Thus, it is extremely important for us to be able to distinguish between an association due to a causal relationship and an association due to confounding (noncausal).

Let us look at another example. For many years it has been known that cigarette smoking by pregnant women is associated with low birth weight in their infants. As seen in Figure 14-7, the effect is not just the result of the birth of a few low-birth-weight babies in this group of women. Rather, the entire weight distribution curve is shifted to the left in the babies born to smokers. The reduction in birth weight is also not a result of shorter pregnancies. The babies of smokers are smaller than those of nonsmokers at each gestational age (Fig. 14-8). A dose-response relationship is also seen (Fig. 14-9). The more a woman smokes, the greater her risk of having a low-birth-weight baby. For many years the interpretation of this association was the subject of great controversy. Many believed the association re! ected a causal relation. Others, including a leading statistician, Jacob Yerushalmy, believed the association was due to confounding and was not causal. He wrote as follows:

> *A comparison of smokers and nonsmokers shows that the two differ markedly along many environmental, behavioral and biologic variables. For example, smokers are less likely to use contraceptives and to plan the pregnancy. Smokers are more likely to drink coffee, beer and whiskey and the nonsmoker, tea, milk and wine. The smoker is more likely than the nonsmoker to indulge in these habits to excess. In general, the nonsmokers are revealed to be more moderate than the smokers who are shown to be more extreme and carefree in their mode of life. Some biologic differences are also noted between them: Thus smokers have a higher twinning rate only in whites and their age for menarche is lower than for nonsmokers.*[3]

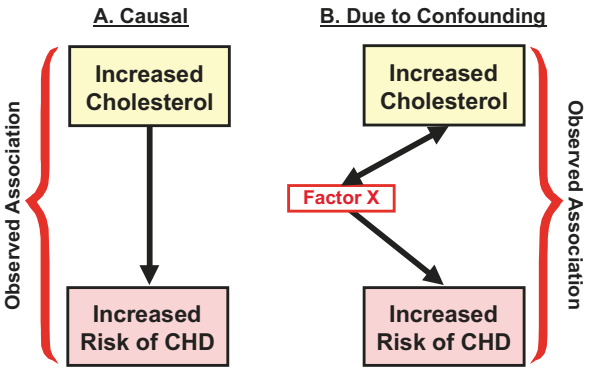


Figure 14-6. Interpreting an observed association between increased cholesterol level and increased risk of coronary heart disease (CHD).

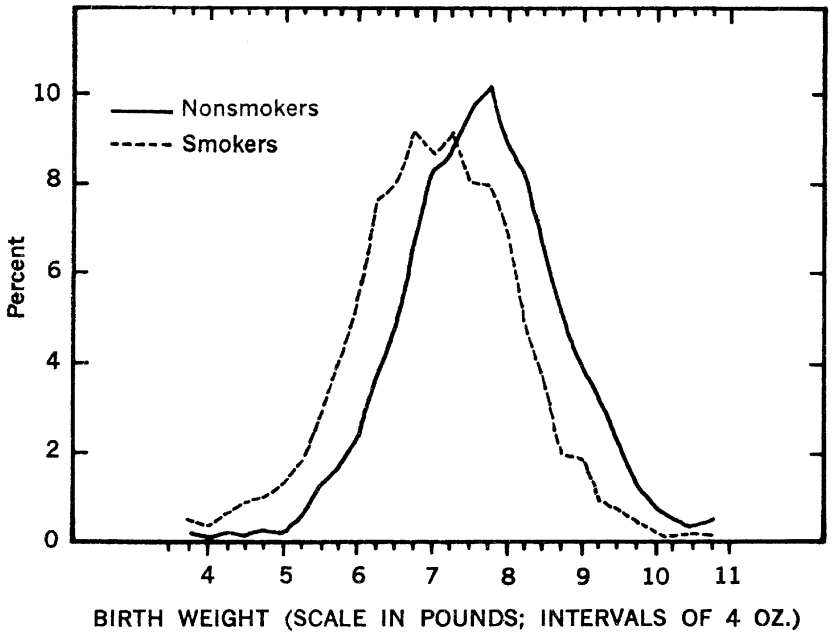


Figure 14-7. Percentage distribution by birth weight of infants of mothers who did not smoke during pregnancy and of those mothers who smoked 1 pack of cigarettes or more per day. (From U.S. Department of Health, Education, and Welfare: The Health Consequences of Smoking. Washington, DC, Public Health Service, 1973, p 105.)

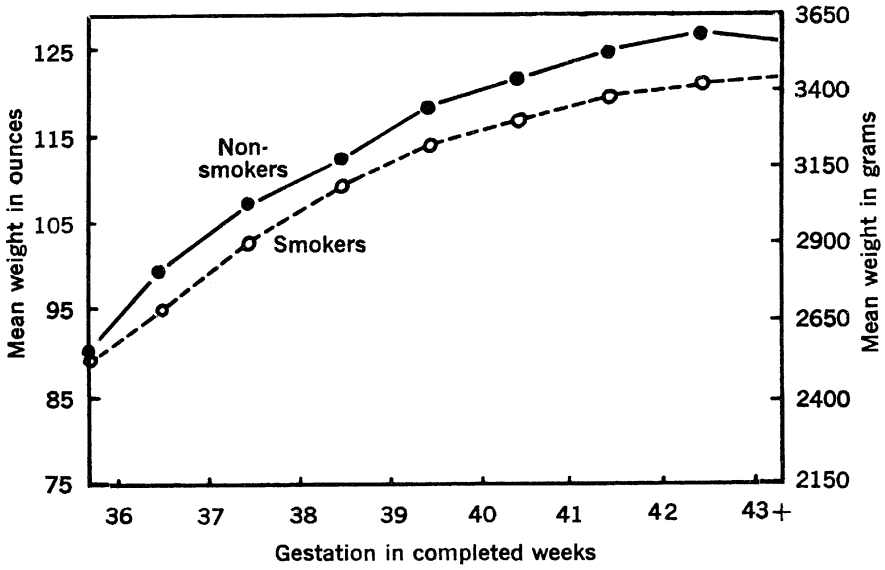


Figure 14-8. Mean birth weight for week of gestation according to maternal smoking habit. (From U.S. Department of Health, Education, and Welfare: The Health Consequences of Smoking. Washington, DC, Public Health Service, 1973, p 104.)

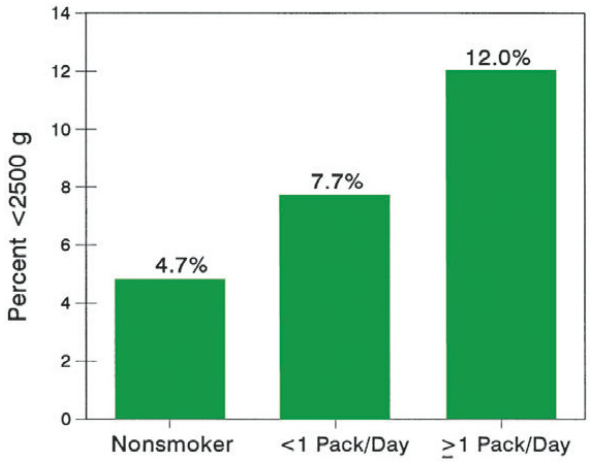


Figure 14-9. Percentage of pregnancies (n = 50,267) with infant weighing less than 2,500 g, by maternal cigarette smoking category. (Redrawn from Ontario Department of Health: Second Report of the Perinatal Mortality Study in Ten University Teaching Hospitals. Toronto, Ontario, Department of Health, Ontario Perinatal Mortality Study Committee, Vol. I, 1967, p 275.)

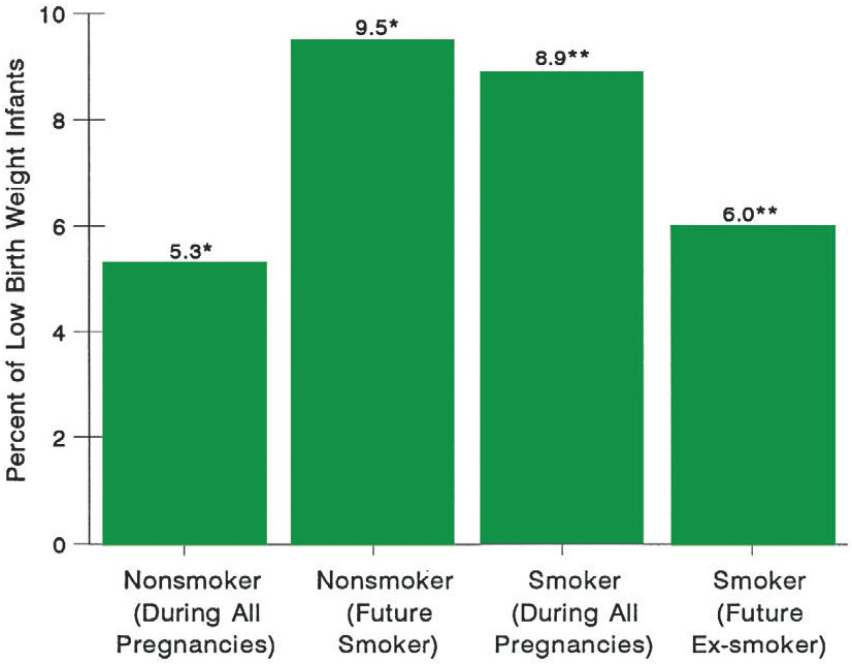


Figure 14-10. Percentage of low-birth-weight infants by smoking status of their mothers (*$P < .01$; **$P < .02$). (Redrawn from Yerushalmy J: Infants with low birth weight born before their mothers started to smoke cigarettes. Am J Obstet Gynecol 112:277–284, 1972.)

In view of these many differences between smokers and nonsmokers, Yerushalmy believed that it was not the *smoking* that caused the low birth weight, but rather that the low weight was attributable to *other characteristics of the smokers*. It is interesting to examine a study that Yerushalmy carried out to support his position at the time (Fig. 14-10).[3]

Yerushalmy examined the results of one pregnancy (the study pregnancy) in a population of women who had had several pregnancies. The rate of low-birth-weight babies in the study pregnancy was 5.3% for women who were nonsmokers in *all* of their pregnancies. However, if they were smokers in all of their pregnancies, the rate of low birth weight in the study pregnancy was almost 9%. When he examined pregnancies of women who were nonsmokers during the study pregnancy, but who later became smokers, he found that their rate of low-birth-weight babies was about equal to that of women who smoked in all pregnancies. When he examined pregnancies of women who were smokers in the study pregnancy, but who subsequently stopped smoking, he found that their rate of low birth weight in the study pregnancy was similar to that of women who were nonsmokers in all of their pregnancies.

On the basis of these data, Yerushalmy came to the conclusion that it was not the smoking but rather some characteristic of the smoker that caused the low birth weight. Today, however, it is virtually universally accepted that smoking is a cause of low birth weight. The causal nature of this relation has also been demonstrated in randomized trials that have reduced the frequency of low birth weight by initiating programs for smoking cessation in pregnant women. Although this issue has now largely been resolved, it is illuminating to review both the controversy and the study, as they exemplify the reasoning that is necessary in trying to distinguish causal from noncausal interpretations of observed associations.

TYPES OF CAUSAL RELATIONSHIPS

A causal pathway can be either *direct* or *indirect* (Fig. 14-11). In *direct* causation, a factor directly

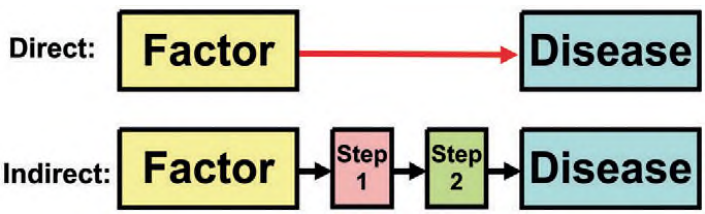


Figure 14-11. Direct versus indirect causes of disease.

Figure 14-12. Types of causal relationships: I. Factor A is both necessary and suf! cient.

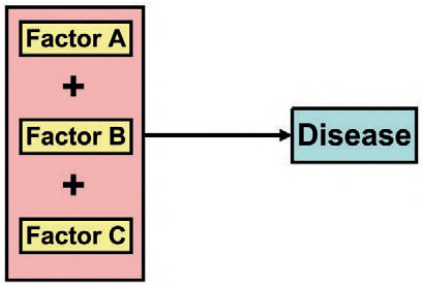


Figure 14-13. Types of causal relationships: II. Each factor is necessary, but not suf! cient.

causes a disease without any intermediate step. In *indirect* causation, a factor causes a disease, but only through an intermediate step or steps. In human biology, intermediate steps are virtually always present in any causal process.

If a relationship is causal, four types of causal relationships are possible: (1) necessary and suf! cient; (2) necessary, but not suf! cient; (3) suf! cient, but not necessary; and (4) neither suf! cient nor necessary.

Necessary and Sufficient

In the ! rst type of causal relationship, a factor is both necessary and suf! cient for producing the disease. Without that factor, the disease never develops (the factor is necessary), and in the presence of that factor, the disease always develops (the factor is suf! cient) (Fig. 14-12). This situation rarely if ever occurs. For example, in most infectious diseases, a number of people are exposed, some of whom will manifest the disease and others who will not. Members of households of a person with tuberculosis do not uniformly acquire the disease from the index case. If the exposure dose is assumed to be the same, there are likely differences in immune status, genetic susceptibility, or other characteristics that determine who develops the disease and who does not. A one-to-one relationship of exposure to disease, which is a consequence of a necessary and suf! cient relationship, rarely if ever occurs.

Necessary, But Not Sufficient

In another model, each factor is necessary, but not, in itself, suf! cient to cause the disease (Fig. 14-13). Thus, multiple factors are required, often in a speci! c temporal sequence. For example,

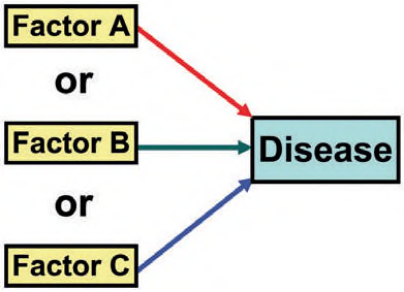


Figure 14-14. Types of causal relationships: III. Each factor is suf! cient, but not necessary.

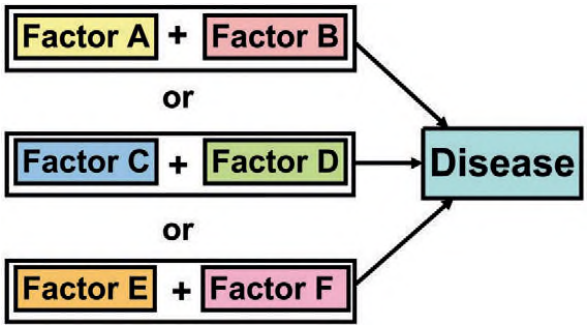


Figure 14-15. Types of causal relationships: IV. Each factor is neither suf! cient nor necessary.

carcinogenesis is considered to be a multistage process involving both initiation and promotion. For cancer to result, a promoter must act after an initiator has acted. Action of an initiator or a promoter alone will not produce a cancer.

Again, in tuberculosis, the tubercle bacillus is clearly a necessary factor, even though its presence may not be suf! cient to produce the disease in every infected individual.

Sufficient, But Not Necessary

In this model, the factor alone can produce the disease, but so can other factors that are acting alone (Fig. 14-14). Thus, either radiation exposure or benzene exposure can each produce leukemia without the presence of the other. Even in this situation, however, cancer does not develop in everyone who has experienced radiation or benzene exposure, so although both factors are not needed, other cofactors probably are. Thus, the criterion of *suf! cient* is rarely met by a single factor.

Neither Sufficient Nor Necessary

In the fourth model, a factor, by itself, is neither suf! cient nor necessary to produce disease (Fig. 14-15). This is a more complex model, which probably most accurately represents the causal relationships that operate in most chronic diseases.

EVIDENCE FOR A CAUSAL RELATIONSHIP

Many years ago, when the major disease problems faced by man were infectious in origin, the question arose as to what evidence would be necessary to prove that an organism causes a disease. In 1840, Henle proposed postulates for causation that were expanded by Koch in the 1880s.[4] The postulates for causation were as follows:

1. The organism is *always* found with the disease.
2. The organism is *not* found with any other disease.
3. The organism, when isolated from one who has the disease, and cultured through several generations, produces the disease (in experimental animals).

Koch added that "Even when an infectious disease cannot be transmitted to animals, the 'regular' and 'exclusive' presence of the organism [postulates 1 and 2] proves a causal relationship."[4]

These postulates, though not perfect, proved very useful for infectious diseases. However, as apparently noninfectious diseases assumed increasing importance toward the middle of the 20th century, the issue arose as to what would represent strong evidence of causation in diseases that were generally not of infectious origin. In such disease there was no organism that could be cultured and grown in animals. Speci! cally, as attention was directed to a possible relationship between smoking and lung cancer, the U.S. Surgeon General appointed an expert committee to review the evidence. The committee developed a set of guidelines,[5] which have been revised over the years. The next few pages present a modi! ed list of these guidelines (Table 14-1) with some brief comments.

TABLE 14-1. **Guidelines for Judging Whether an Observed Association Is Causal**

1. Temporal relationship
2. Strength of the association
3. Dose-response relationship
4. Replication of the ! ndings
5. Biologic plausibility
6. Consideration of alternate explanations
7. Cessation of exposure
8. Consistency with other knowledge
9. Speci! city of the association

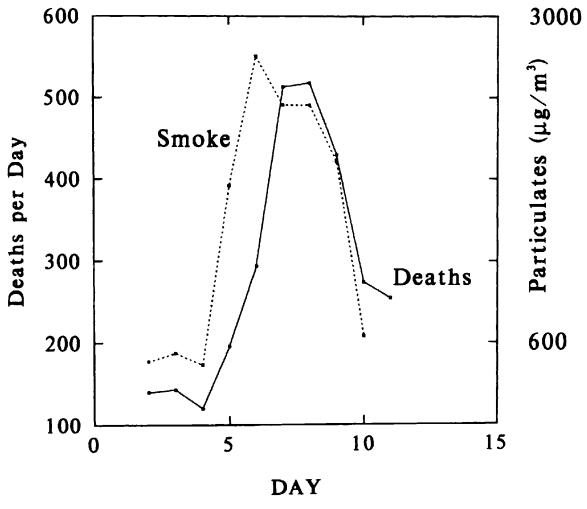


Figure 14-16. The mean concentration of airborne particles (µg/m³) from the four inner monitoring stations in London and the count of daily deaths in the London Administrative County during the beginning of December 1952. (From Schwartz J: Air pollution and daily mortality: A review and meta analysis. Environ Res 64:36–52, 1994.)

GUIDELINES FOR JUDGING WHETHER AN OBSERVED ASSOCIATION IS CAUSAL

1. Temporal Relationship. It is clear that if a factor is believed to be the cause of a disease, exposure to the factor must have occurred before the disease developed. Figure 14-16 shows the number of deaths per day and the mean concentration of airborne particles in London in early December 1952.[6] The pattern of a rise in particle concentration followed by a rise in mortality and a subsequent decline in particle concentration followed by a decline in mortality strongly supported the increase in mortality being due to the increase in air pollution. This example demonstrates the use of ecologic data for exploring a temporal relationship. Further investigation revealed that the increased mortality consisted almost entirely of respiratory and cardiovascular deaths and was highest in the elderly.

It is often easier to establish a temporal relationship in a prospective cohort study than in a case-control study or a retrospective cohort study. In the last two types of studies, exposure information may need to be obtained or re-created from past records and the timing may therefore be imprecise.

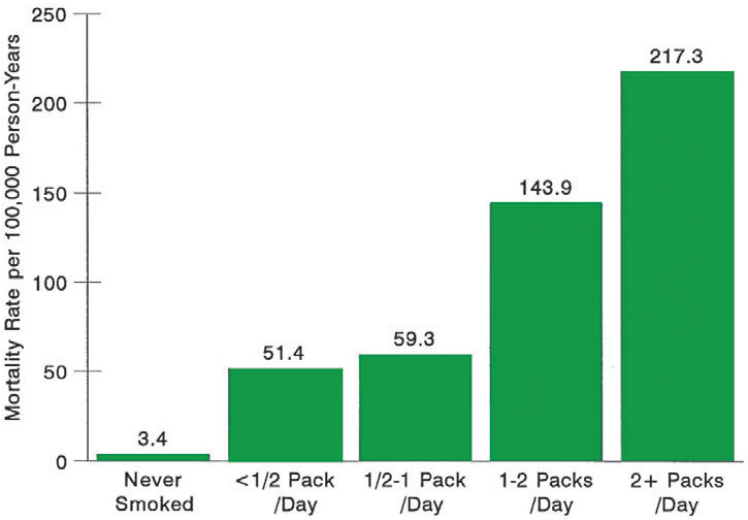


Figure 14-17. Age-standardized death rates due to well-established cases of bronchogenic carcinoma (exclusive of adenocarcinoma) by current amount of smoking. (Adapted from Hammond EC, Horn D: Smoking and death rates: Report on 44 months of follow-up of 187,783 men: II. Death rates by cause. JAMA 166:1294–1508, 1958. Copyright 1958, American Medical Association.)

The temporal relationship of exposure and disease is important not only for clarifying the order in which the two occur but also in regard to the length of the interval between exposure and disease. For example, asbestos has been clearly linked to increased risk of lung cancer, but the latent period between the exposure and the appearance of lung cancer is at least 15 to 20 years. Therefore, if, for example, lung cancer develops after only 3 years since the asbestos exposure, it is safe to conclude that the lung cancer was not a result of this exposure.

2. Strength of the Association. The strength of the association is measured by the relative risk (or odds ratio). The stronger the association, the more likely it is that the relation is causal.

3. Dose-Response Relationship. As the dose of exposure increases, the risk of disease also increases. Figure 14-17 shows an example of the dose-response relationship for cigarette smoking and lung cancer. If a dose-response relationship is present, it is strong evidence for a causal relationship. However, the absence of a dose-response relationship does not necessarily rule out a causal relationship. In some cases in which a threshold may exist, no disease may develop up to a certain level of exposure (a threshold); above this level, disease may develop.

4. Replication of the Findings. If the relationship is causal, we would expect to find it consistently in different studies and in different populations. Replication of findings is particularly important in epidemiology. If an association is observed, we would also expect it to be seen consistently within subgroups of the population and in different populations, unless there is a clear reason to expect different results.

5. Biologic Plausibility. Biologic plausibility refers to coherence with the current body of biologic knowledge. Examples may be cited to demonstrate that epidemiologic observations have sometimes preceded biologic knowledge. Thus, as discussed in an earlier chapter, Gregg's observations on rubella and congenital cataracts preceded any knowledge of teratogenic viruses. Similarly, the implication of high oxygen concentration in the causation of retrolental fibroplasia, a form of blindness that occurs in premature infants, preceded any biologic knowledge supporting such a relationship. Nevertheless, we seek consistency of the epidemiologic findings with existing biologic knowledge, and when this is not the case, interpreting the meaning of the observed association may be difficult. We may then be more demanding in our requirements about the size and significance of any differences observed and in having the study replicated by other investigators in other populations.

6. Consideration of Alternate Explanations. We have discussed the problem in interpreting an observed association in regard to whether a relationship is causal or is the result of confounding. In judging whether a reported association is causal, the extent to which the investigators have taken other possible explanations into account and the extent to which they have ruled out such explanations are important considerations.

7. Cessation of Exposure. If a factor is a cause of a disease, we would expect the risk of the disease to decline when exposure to the factor is reduced

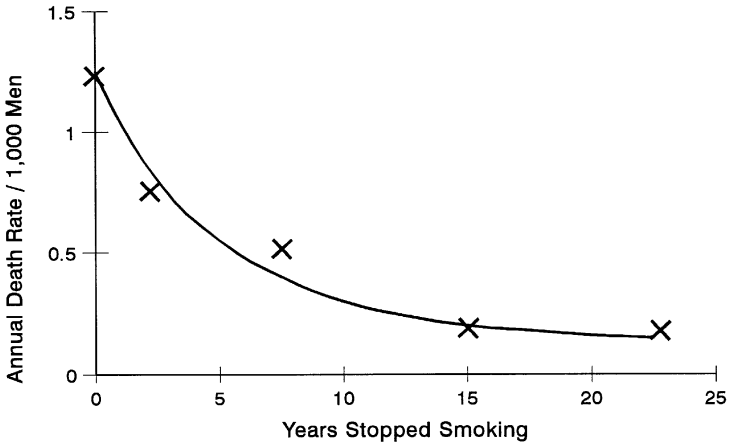


Figure 14-18. Effects of terminating exposure: lung cancer death rates, standardized for age and amount smoked, among men continuing to smoke cigarettes and men who gave up smoking for different periods. The corresponding rate for nonsmokers was 0.07 per 1,000. (Adapted from Doll R, Hill AB: Mortality in relation to smoking: Ten years' observations of British doctors. BMJ 1:1399–1410, 1964.)

or eliminated. Figure 14-18 shows such data for cigarette smoking and lung cancer.

Eosinophilia-myalgia syndrome (EMS) reached epidemic proportions in 1989. Characterized by severe muscle pain and a high blood eosinophil count, the syndrome was found to be associated with manufactured preparations of L-tryptophan. In November 1989, a nationwide recall by the Food and Drug Administration of over-the-counter preparations of L-tryptophan was followed by dramatic reductions in numbers of cases of EMS reported each month (Fig. 14-19). This is another example of a reduction in incidence being related to cessation of exposure, which adds to the strength of the causal inference regarding the exposure.

When cessation data are available, they provide helpful supporting evidence for a causal association. However, in certain cases, the pathogenic process may have been irreversibly initiated, and the disease occurrence may have been determined by the time the exposure is removed. Emphysema is not reversed with cessation of smoking, but its progression is reduced.

8. Consistency with Other Knowledge. If a relationship is causal, we would expect the ! ndings to be consistent with other data. For example, Figure 14-20 shows data regarding lung cancer rates in men and women and cigarette smoking in men and women.

We see a consistent direction in the curves, with the increase in lung cancer rates following the increase in cigarette sales in both men and women. These data are consistent with what we would expect if the relationship between smoking and lung cancer is established as a causal one. Although the absence of such consistency would not completely rule out this hypothesis, if we observed rising lung cancer rates after a period of declining

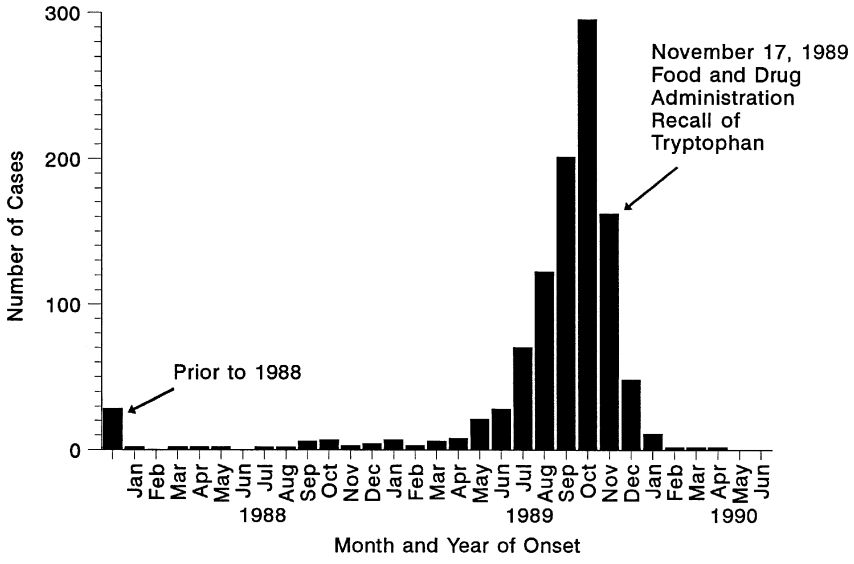


Figure 14-19. Reported dates of illness onset by month and year for cases of eosinophilia-myalgia syndrome, as reported to the Centers for Disease Control and Prevention, Atlanta, as of July 10, 1990. (Adapted from Swygert LA, Maes EF, Sewell LE, et al: Eosinophilia-myalgia syndrome: Results of national surveillance. JAMA 264:1698–1703, 1990. Copyright 1990, American Medical Association.)

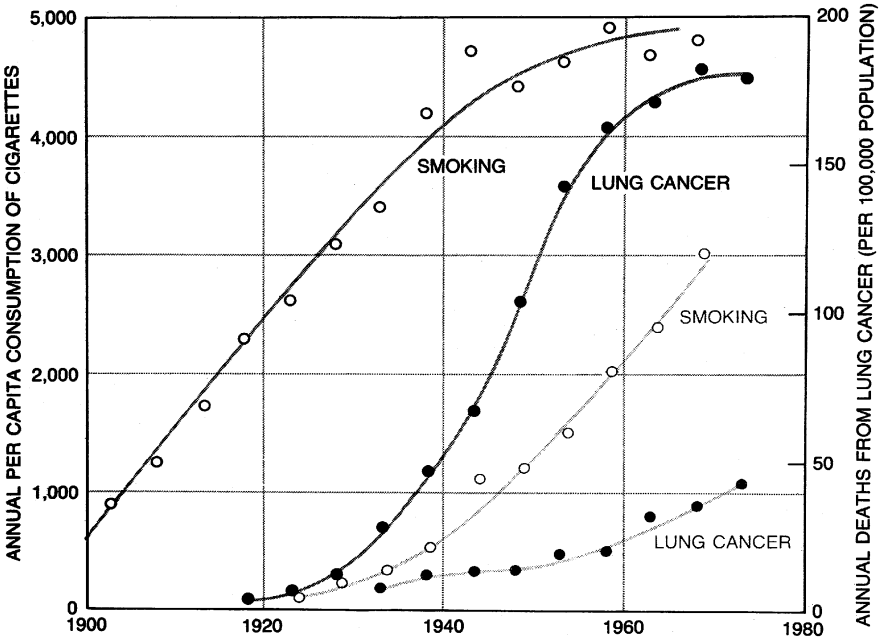


Figure 14-20. Parallel trends between cigarette consumption and lung cancer in men (two curves on left) and in women (two curves on right), in England and Wales. (From Cairns J: The cancer problem. Sci Am 233:64–72, 77–78, 1975.)

cigarette sales, for example, we would need to explain how this observation could be consistent with a causal hypothesis.

9. Specificity of the Association. An association is speci! c when a certain exposure is associated with only one disease; this is the weakest of all the guidelines and should probably be deleted from the list. Cigarette manufacturers have pointed out that the diseases attributed to cigarette smoking do not meet the requirements of this guideline, because cigarette smoking has been linked to lung cancer, pancreatic cancer, bladder cancer, heart disease, emphysema, and other conditions.

The possibility of such multiple effects from a single factor is not, in fact, surprising: regardless of the tissue that comprises them, all cells have common characteristics, including DNA, RNA, and various subcellular structures, so a single agent could have effects in multiple tissues. Furthermore, cigarettes are not a single factor but constitute a mixture of a large number of compounds; consequently, a large number of effects might be anticipated.

When speci! city of an association is found, it provides additional support for a causal inference. However, as with a dose-response relationship, absence of speci! city in no way negates a causal relationship.

Any conclusion that an observed association is causal is greatly strengthened when different types of evidence from multiple sources support such reasoning. Thus, it is not so much a count of the number of guidelines present that is relevant to causal inference but rather an *assessment of the total pattern of evidence observed* that may be consistent with one or more of the guidelines. Sir Austin Bradford Hill eloquently expressed this sentiment in an essay written in 1965:

> *Here then are nine different viewpoints [guidelines] from all of which we should study association before we cry causation. What I do not believe—and this has been suggested—that we can usefully lay down some hard-and-fast rules of evidence that must be obeyed before we can accept cause and effect. None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a sine qua non. What they can do, with greater or less strength, is to help us to make up our minds on the fundamental question—is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?*[7]

DERIVING CAUSAL INFERENCES: TWO EXAMPLES

Peptic Ulcers and Gastric Cancer in Relation to Infection with *Helicobacter pylori*

Although the preceding guidelines do not permit a quantitative estimation of whether or not an

association is causal, they can nevertheless be very helpful, as seen in the following examples:

Until the 1980s, the major causes of peptic ulcer disease were considered to be stress and lifestyle factors, including smoking. Peptic ulcer disease had long been attributed to the effects of gastric acid. Susceptibility to gastric acid had been linked to cigarette smoking, alcohol consumption, and use of nonsteroidal anti-in! ammatory agents. Therapy was primarily directed at inhibiting acid secretion and protecting mucosal surfaces from acid. Although these therapies helped healing, relapses were common.

In 1984, Australian physicians Drs. Barry J. Marshall and J. Robin Warren reported that they had observed small curved bacteria colonizing the lower part of the stomach in patients with gastritis and peptic ulcers.[8] After several attempts, Marshall succeeded in cultivating a hitherto unknown bacterial species (later named *Helicobacter pylori*) from several of these biopsies (Fig. 14-21). Together they found that the organism was present in almost all patients with gastric in! ammation or peptic ulcer. Many of these patients had biopsies performed which showed evidence of in! ammation present in the gastric mucosa close to where the bacteria were seen. Based on these results, they proposed that *Helicobacter pylori* is involved in the etiology of these diseases. It was subsequently shown that the ulcer was often not cured until *Helicobacter pylori* had been eliminated.

It is now ! rmly established that *Helicobacter pylori* causes more than 90% of duodenal ulcers and up to 80% of gastric ulcers. The link between *Helicobacter pylori* infection and subsequent gastritis and peptic ulcer disease has been established through studies of human volunteers, antibiotic treatment studies, and epidemiological studies. Thus, many of the study designs discussed in previous chapters and many of the guidelines for causal inferences discussed earlier in this chapter were involved in elucidating the role of *Helicobacter pylori* in peptic ulcer and gastritis. In 2005, the Nobel Prize for Physiology or Medicine was shared by Drs. Marshall and Warren, "for their discovery of the bacterium *Helicobacter pylori* and its role in gastritis and peptic ulcer disease."

Table 14-2 categorizes this evidence according to several of the guidelines for causation just discussed. Thus, as seen here, the guidelines can be extremely helpful in characterizing the evidence supporting a causal relationship.

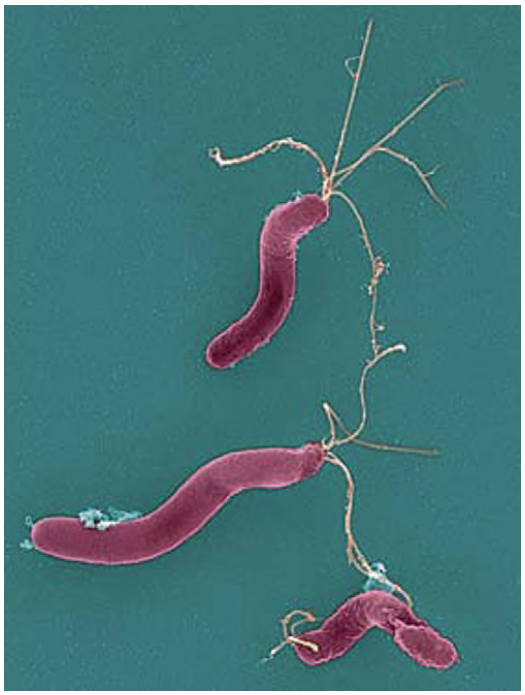


Figure 14-21. *Helicobacter pylori* [Photograph]. (Encyclopædia Britannica Online. http://www.britannica.com/EBchecked/topic/450889/peptic-ulcer?overlay=true&assemblyId=94921. Accessed August 15, 2013.)

Increasing evidence now also supports the association of *Helicobacter pylori* infection and the development of gastric cancer. Uemura and coworkers[9] prospectively studied 1,526 Japanese patients who had duodenal or gastric ulcers, gastric hyperplasia, or nonulcer hyperplasia. Of this group, 1,246 had *Helicobacter pylori* infection and 280 did not. The mean follow-up period was 7.8 years. Gastric cancers developed in 36 (2.9%) of the infected patients, but in none of the noninfected patients. Individuals who carry antibodies to *Helicobacter pylori* may have a 2 to 3 times higher risk of stomach cancer than those who do not (Fig. 14-22). The risk of stomach cancer also appears to be related to the type of strain of *Helicobacter pylori* which is infecting a person. Evidence is accumulating to support the idea that therapy against *Helicobacter pylori* may prevent gastric cancer. In the future, gastric cancer may come to be viewed as a largely preventable cancer of infectious origin.

Age of Onset of Alcohol Use and Lifetime Alcohol Abuse

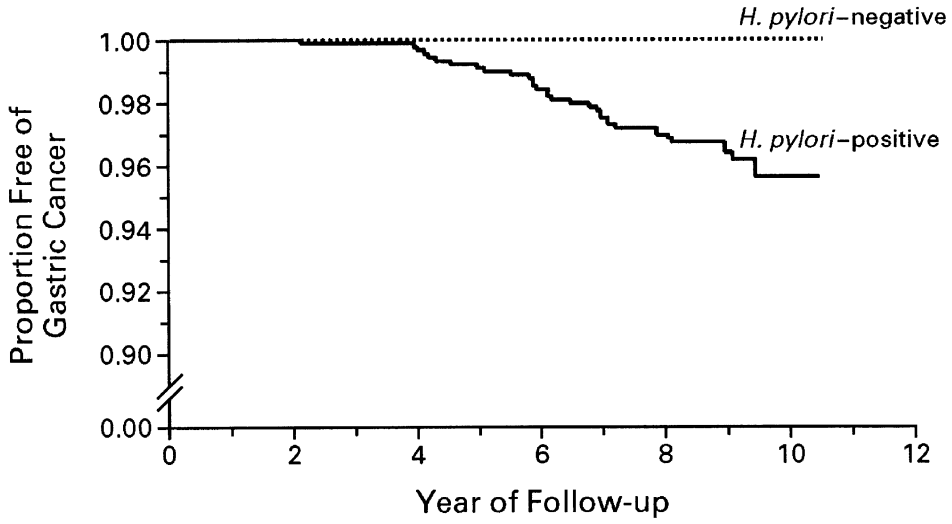
In 1997, Grant and Dawson[10] reported data on the relationship of age at ! rst use of alcohol and prevalence of lifetime alcohol dependence and abuse.

TABLE 14-2. **Assessment of the Evidence Suggesting *Helicobacter pylori* as a Causative Agent of Duodenal Ulcers**

1. **Temporal relationship.**
   - *Helicobacter pylori* is clearly linked to chronic gastritis. About 11% of chronic gastritis patients will go on to have duodenal ulcers over a 10-year period.
   - In one study of 454 patients who underwent endoscopy 10 years earlier, 34 of 321 patients who had been positive for *Helicobacter pylori* (11%) had duodenal ulcer compared with 1 of 133 *Helicobacter pylori*–negative patients (0.8%).
2. **Strength of the association.**
   - *Helicobacter pylori* is found in at least 90% of patients with duodenal ulcer. In at least one population reported to lack duodenal ulcers, a northern Australian aboriginal tribe that is isolated from other people, it has never been found.
3. **Dose-response relationship.**
   - Density of *Helicobacter pylori* per square millimeter of gastric mucosa is higher in patients with duodenal ulcer than in patients without duodenal ulcer. Also see item 2 above.
4. **Replication of the findings.**
   - Many of the observations regarding *Helicobacter pylori* have been replicated repeatedly.
5. **Biologic plausibility.**
   - Although originally it was dif! cult to envision a bacterium that infects the stomach antrum causing ulcers in the duodenum, it is now recognized that *Helicobacter pylori* has binding sites on antral cells and can follow these cells into the duodenum.
   - *Helicobacter pylori* also induces mediators of in! ammation.
   - *Helicobacter pylori*–infected mucosa is weakened and is susceptible to the damaging effects of acid.
6. **Consideration of alternate explanations.**
   - Data suggest that smoking can increase the risk of duodenal ulcer in *Helicobacter pylori*–infected patients but is not a risk factor in patients in whom *Helicobacter pylori* has been eradicated.
7. **Cessation of exposure.**
   - Eradication of *Helicobacter pylori* heals duodenal ulcers at the same rate as histamine receptor antagonists.
   - Long-term ulcer recurrence rates were zero after *Helicobacter pylori* was eradicated using triple-antimicrobial therapy, compared with a 60% to 80% relapse rate often found in patients with duodenal ulcers treated with histamine receptor antagonists.
8. **Consistency with other knowledge.**
   - Prevalence of *Helicobacter pylori* infection is the same in men as in women. The incidence of duodenal ulcer, which in earlier years was believed to be higher in men than in women, has been equal in recent years.
   - The prevalence of ulcer disease is believed to have peaked in the latter part of the 19th century, and the prevalence of *Helicobacter pylori* may have been much higher at that time because of poor living conditions. This reasoning is also based on observations today that the prevalence of *Helicobacter pylori* is much higher in developing countries.
9. **Specificity of the association.**
   - Prevalence of *Helicobacter pylori* in patients with duodenal ulcers is 90% to 100%. However, it is found in some patients with gastric ulcer and even in asymptomatic individuals.

Data from Megraud F, Lamouliatte H: *Helicobacter pylori* and duodenal ulcer: Evidence suggesting causation. Dig Dis Sci 37:769–772, 1992; and DeCross AJ, Marshall BJ: The role of *Helicobacter pylori* in acid-peptic disease. Am J Med Sci 306: 381–392, 1993.

They analyzed data from 27,616 current and former drinkers who were interviewed as part of the 1992 National Longitudinal Alcohol Epidemiologic Survey. The rates of lifetime *dependence* decreased from more than 40% among individuals who began drinking at age 14 years or younger to about 10% among those who started drinking at age 20 years or older (Fig. 14-23). The con! guration of the curve in Figure 14-23 suggests a dose-response relationship as has been observed for longer duration of smoking associated with increased risk of lung cancer. However, the data may also point to a period of particularly high susceptibility, namely, that the period of preadolescence and early adolescence is a period of increased risk for developing a disorder of alcohol use. Therefore, interventions



No. at Risk						
H. pylori–negative	280	272	251	245	213	57
H. pylori–positive	1246	1219	1086	907	782	258

Figure 14-22. Kaplan-Meier analysis of the proportion of *Helicobacter pylori*–positive and *Helicobacter pylori*–negative patients who remained free of gastric cancer. During follow-up, gastric cancer developed in 36 of the 1,246 *H. pylori*–infected patients (2.9%), but in none of the 280 uninfected patients ($P < .001$). (From Uemura N, Okamoto S, Yamomoto S, et al: *Helicobacter pylori* infection and the development of gastric cancer. N Engl J Med 345:784–789, 2001.)

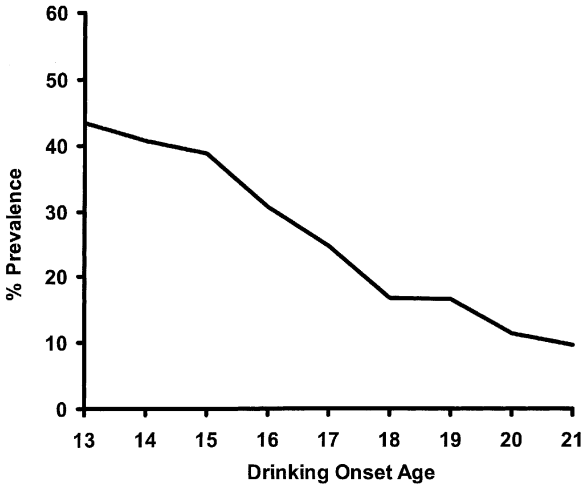


Figure 14-23. Relation of age of onset of alcohol use to prevalence of lifetime alcohol abuse. (Adapted from Grant BF, Dawson DA: Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. J Subst Abuse 9:103–110, 1997.)

should be targeted to this group in the hope of delaying drinking onset. However, adopting such an approach assumes that the relationship between early onset of drinking and subsequent lifetime abuse is a causal one, so that delaying age at onset of drinking would reduce the risk of lifetime alcohol dependence. Another possible explanation is that those who are destined for lifetime alcohol dependence tend to begin drinking earlier, but that the earlier age at drinking onset is not necessarily a cause of the later dependence. Further research is therefore needed to explain the intriguing association that has been observed. We shall return to this example in Chapter 16.

MODIFICATIONS OF THE GUIDELINES FOR CAUSAL INFERENCES

In 1986, the U.S. Public Health Service brought together a group of 19 experts to examine the scienti! c basis of the content of prenatal care and to answer the question: Which measures implemented during prenatal care have actually been demonstrated to be associated with improved outcome? The panel's report was issued in 1989 and served as the basis of a comprehensive report.[11] As the panel began its deliberations, it became clear that questions of causation were at the heart of the panel's task, and that guidelines were needed for assessing the relationship of prenatal measures to health outcomes. A subcommittee reviewed the current guidelines (just enumerated in the preceding text) and de! ned a process for using evidence that includes (1) categorization of the evidence by the

TABLE 14-3. **The Process for Using the Evidence in Developing Recommendations on the Effectiveness of Prenatal Interventions**

Stage I: Categorizing the Evidence by the Quality of Its Source. (*In each category, studies are listed in descending order of quality.*)

1. Trials (planned interventions with contemporaneous assignment of treatment and nontreatment)
   a. Randomized, double-blind, placebo-controlled with sufficient power appropriately analyzed.
   b. Randomized, but blindness not achieved.
   c. Nonrandomized trials with good control of confounding, that are well conducted in other respects.
   d. Randomized, but with deficiencies in execution or analysis (insufficient power, major losses to follow-up, suspect randomization, analysis with exclusions).
   e. Nonrandomized trials with deficiencies in execution or analysis.
2. Cohort or case-control studies
   a. Hypothesis specified before analysis, good data, confounders accounted for.
   b. As above, but hypothesis not specified before analysis.
   c. Post hoc, with problem(s) in the data or the analysis.
3. Time-series studies
   a. Analyses that take confounding into account.
   b. Analyses that do not consider confounding.
4. Case-series studies: Series of case reports without any specific comparison group

Among other issues that must be considered in reviewing the evidence are the precision of definition of the outcome being measured, the degree to which the study methodology has been described, adequacy of the sample size, and the degree to which characteristics of the population studied and of the intervention being evaluated have been described.

A study can be well designed and carried out in an exemplary fashion (internal validity), but if the population studied is an unusual or highly selected one, the results may not be generalizable (external validity).

Stage II: Guidelines for Evaluating the Evidence of a Causal Relationship. (*In each category, studies are listed in descending priority order.*)

1. Major criteria
   a. Temporal relationship: An intervention can be considered evidence of a reduction in risk of disease or abnormality only if the intervention was applied before the time the disease or abnormality would have developed.
   b. Biological plausibility: A biologically plausible mechanism should be able to explain why such a relationship would be expected to occur.
   c. Consistency: Single studies are rarely definitive. Study findings that are replicated in different populations and by different investigators carry more weight than those that are not. If the findings of studies are inconsistent, the inconsistency must be explained.
   d. Alternative explanations (confounding): The extent to which alternative explanations have been explored is an important criterion in judging causality.
2. Other considerations
   a. Dose-response relationship: If a factor is indeed the cause of a disease, usually (but not invariably) the greater the exposure to the factor, the greater the risk of the disease. Such a dose-response relationship may not always be seen because many important biologic relationships are dichotomous, and reach a threshold level for observed effects.
   b. Strength of the association: The strength of the association is usually measured by the extent to which the relative risk or odds depart from unity, either above 1 (in the case of disease-causing exposures) or below 1 (in the case of preventive interventions).
   c. Cessation effects: If an intervention has a beneficial effect, then the benefit should cease when it is removed from a population (unless carryover effect is operant).

Adapted from Gordis L, Kleinman JC, Klerman LV, et al: Criteria for evaluating evidence regarding the effectiveness of prenatal interventions. In Merkatz IR, Thompson JE (eds): New Perspectives on Prenatal Care. New York, Elsevier, 1990, pp 31–38.

quality of its sources, and (2) evaluation of the evidence of a causal relationship using standardized guidelines.[12] These recommendations are excerpted in Table 14-3. Although these modi! ed guidelines clearly use the original components, they establish reasonable priorities in weighting them. They thus de! ne an approach for looking at causation that may have applicability far beyond questions of the effectiveness of prenatal measures.

A similar approach, ranking studies by the quality of the study and its evidence, is used by the U.S. Preventive Services Task Force, which is responsible for developing clinical practice guidelines for prevention and screening (Table 14-4).[13] The Task Force is an independent committee of experts supported by the U.S. Government. Members include experts in primary care, prevention, evidence-based medicine, and research methods. Various clinical areas and experience in preventive medicine, public health, and health policy are also represented.

For each topic the Task Force considers, it de! nes the questions that need to be addressed and identi! es and retrieves the relevant evidence. The quality of each individual study is assessed after which the strength of the totality of available evidence is judged. Estimates are made of the balance of bene! ts and harms. This balance is expressed as the net bene! t (the difference between bene! ts and harms). The Task Force prepares recommendations for preventive interventions based on these considerations.

Figure 14-24 shows a generic example of the analytic plan which is prepared by the Task Force as a framework for evaluating the evidence for a screening program. The straight arrows show possible pathways of bene! t, and the blue curved arrows show possible adverse effects relating to different stages. The primary question (question 1 in the ! gure) is generally one of whether screening is effective in reducing the risk of an adverse outcome such as mortality and if so, to what extent.

Generally, few if any studies have examined this overarching question so that the deliberations of the Task Force often deal with the different steps or linkages that comprise this overall pathway. The

TABLE 14-4. U.S. Preventive Services Task Force Levels of Certainty* Regarding Net Benefit

HIGH	The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.
MODERATE	The available evidence is suf! cient to determine the effects of the preventive service on health outcomes, but con! dence in the estimate is constrained by such factors as: • The number, size, or quality of individual studies. • Inconsistency of ! ndings across individual studies. • Limited generalizability of ! ndings to routine primary care practice. • Lack of coherence in the chain of evidence. As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion.
LOW	The available evidence is insuf! cient to assess effects on health outcomes. Evidence is insuf! cient because of: • The limited number or size of studies. • Important ! aws in study design or methods. • Inconsistency of ! ndings across individual studies. • Gaps in the chain of evidence. • Findings not generalizable to routine primary care practice. • A lack of information on important health outcomes. More information may allow an estimation of effects on health outcomes.

*The USPSTF de! nes certainty as "likelihood that the USPSTF assessment of the net bene! t of a preventive service is correct." The net bene! t is de! ned as bene! t minus harm of the preventive service as implemented in a general, primary care population. The USPSTF assigns a certainty level based on the nature of the overall evidence available to assess the net bene! t of a preventive service.

From the U.S. Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF, July 2008. http://www.uspreventiveservicestaskforce.org/uspstf08/methods/procmanual.htm/. Accessed August 15, 2013.

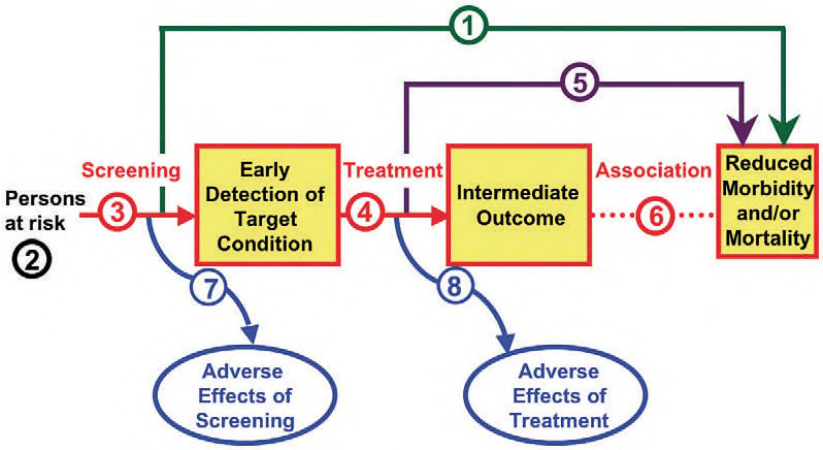


Figure 14-24. Generic analytic framework for screening topics used by the U.S. Preventive Services Task Force. Numbers refer to key questions in the ! gure. (1) Does screening for X reduce morbidity and/or mortality? (2) Can a group at high risk for X be identi- ! ed on clinical grounds? (3) Are there accurate (i.e., sensitive and speci! c) screening tests available? (4) Are treatments available that make a difference in intermediate outcomes when the disease is caught early, or detected by screening? (5) Are treatments available that make a difference in morbidity or mortality when the disease is caught early, or detected by screening? (6) How strong is the association between the intermediate outcomes and patient outcomes? (7) What are the harms of the screening test? (8) What are the harms of the treatment? (Adapted from U.S. Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF, July 2008. http://www.uspreventiveservicestaskforce.org/uspstf08/methods/procmanual.htm. Accessed August 15, 2013.)

purple arrow in the ! gure (step 5) shows the relation of treatment to outcome. Red arrows in the ! gure, steps 3, 4, and 6, show individual components of question 1. These assessments generally depend on a review of relevant randomized trials in order to prepare a chain of supporting evidence on which to base an answer to question 1. The evidence for each linkage is summarized in the evidence review and then summarized across the different linkages to provide an overall assessment of the supporting evidence for the preventive service being evaluated.

The certainty of net bene! t is graded on a 3-point scale: high, moderate, or low (see Table 14-4). The recommendations of the Task Force are based on a combined consideration of the certainty and the magnitude of the net bene! t as shown in the matrix in Figure 14-25, in which a grading system of A, B, C, D, and I is used. The meaning of each letter grade is explained in Table 14-5.

The work of the Task Force has dealt with screening for many diseases and conditions. Some examples will illustrate the breadth of the Task Force's activities. It has reviewed the evidence for screening for different cancers, for cardiovascular diseases including hypertension, coronary heart disease, and abdominal aortic aneurysm, for infectious diseases, such as gonorrhea, Chlamydia, and hepatitis B and C, and for mental conditions such as dementia, depression, and suicide risk, and screening for glaucoma and for type 2 diabetes. The Task Force has

Certainty of Net Benefit	Magnitude of Net Benefit			
	Substantial	Moderate	Small	Zero/ Negative
High	A	B	C	D
Moderate	B	B	C	D
Low	Insufficient Evidence			

Figure 14-25. Grid used by the U.S Preventive Services Task Force for assessing the certainty of bene! t and the magnitude of net bene! t in determining the grade of its recommendations. (Adapted from U.S. Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF, July 2008. http://www.uspreventiveservicestaskforce.org/uspstf08/methods/procmanual.htm. Accessed August 15, 2013.)

also reviewed the evidence for the effectiveness of counseling for many conditions such as counseling to prevent tobacco use and tobacco-related diseases, counseling to prevent alcohol misuse, counseling to promote a healthy diet, and counseling to promote physical activity. The above issues have been addressed in adults, but childhood conditions have also been reviewed by the Task Force including prevention of dental caries in preschool children, screening for scoliosis in adolescents, newborn hearing screening, screening for visual impairment in children younger than 5 years of age, and screening for obesity in children and adolescents. These and many more evidence reviews and recommendations of the Task Force can be found on the website of the Agency for Health Care Research and

TABLE 14-5. **What the USPSTF Grades Mean and Suggestions for Practice**

Grade	Grade Definitions	Suggestions for Practice
A	The USPSTF recommends the service. There is high certainty that the net benefit is substantial.	Offer/provide this service.
B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.	Offer/provide this service.
C	The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small.	Offer/provide this service only if there are other considerations in support of offering/providing the service in an individual patient.
D	The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.	Discourage the use of this service.
I	The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.	Read "Clinical Considerations" section of USPSTF Recommendation Statement. If offered, patients should understand the uncertainty about the balance of benefits and harms.

From the U.S. Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF, July 2008. http://www.uspreventiveservicestaskforce.org/uspstf08/methods/procmanual.htm. Accessed August 15, 2013.

Quality (www.ahrq.gov). The deliberations and recommendations of the Task Force provide a highly useful model of assessing the strength of the evidence and moving from causal inferences to policy recommendations.

CONCLUSION

Although causal guidelines discussed in this chapter are often referred to as *criteria*, this term does not seem entirely appropriate. Although it may be a desirable goal to place causal inferences on a firm quantitative and structural foundation, at present we generally do not have all the information needed for doing so. The preceding list should therefore be considered to be only guidelines that can be of most value when coupled with reasoned judgment about the entire body of available evidence, in making decisions about causation.

In the next chapter, we address several additional issues that need to be considered in deriving causal inferences from epidemiologic studies.

REFERENCES

1. Cameron WB: Informal Sociology: A Casual Introduction to Sociological Thinking. New York, Random House, 1963, p 13. (This quotation was also attributed to Albert Einstein some years later.)
2. Ochsner A, DeBakey M: Primary pulmonary malignancy. Surg Gynecol Obstet 68:435, 1939.
3. Yerushalmy J: Infants with low birth weight born before their mothers started to smoke cigarettes. Am J Obstet Gynecol 112:277–284, 1972.
4. Evans AS: Causation and Disease: A Chronological Journey. New York, Plenum, 1993, pp 13–39.
5. U.S. Department of Health, Education and Welfare: Smoking and Health: Report of the Advisory Committee to the Surgeon General. Washington, DC, Public Health Service, 1964.
6. Schwartz J: Air pollution and daily mortality: A review and meta analysis. Environ Res 64:36–52, 1994.
7. Hill AB: The environment and disease: Association or causation? Proceedings of the Royal Society of Medicine, 58:295–300, 1965.
8. Marshall BJ, Warren JR: Unidentified curved bacilli in the stomachs of patients with gastritis and peptic ulceration. Lancet 1:1311–1315, 1984.
9. Uemura N, Okamoto S, Yamamoto S, et al: *Helicobacter pylori* infection and the development of gastric cancer. N Engl J Med 345:784–789, 2001.
10. Grant BF, Dawson DA: Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. J Subst Abuse 9:103–110, 1997.

11. Merkatz IR, Thompson JE (eds): New Perspectives on Prenatal Care. New York, Elsevier, 1990.
12. Gordis L, Kleinman JC, Klerman LV, et al: Criteria for evaluating evidence regarding the effectiveness of prenatal interventions. In Merkatz IR, Thompson JE (eds): New Perspectives on Prenatal Care. New York, Elsevier, 1990, pp 31–38.
13. U.S. Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF, July 2008. http://www.uspreventiveservicestaskforce.org/uspstf08/methods/procmanual.htm. Accessed August 15, 2013.

REVIEW QUESTIONS FOR CHAPTER 14

1. In a large case-control study of patients with pancreatic cancer, 17% of the patients were found to be diabetic at the time of diagnosis, compared to 4% of a well-matched control group (matched by age, sex, ethnic group, and several other characteristics) that was examined for diabetes at the same time as the cases were diagnosed. It was concluded that the diabetes played a causal role in the pancreatic cancer. This conclusion:
- **a.** Is correct
- **b.** May be incorrect because there is no control or comparison group
- **c.** May be incorrect because of failure to establish the time sequence between onset of the diabetes and diagnosis of pancreatic cancer
- **d.** May be incorrect because of less complete ascertainment of diabetes in the pancreatic cancer cases
- **e.** May be incorrect because of more complete ascertainment of pancreatic cancer in nondiabetic persons

2. An investigator examined cases of fetal death in 27,000 pregnancies and classi! ed mothers according to whether they had experienced sexual intercourse within 1 month before delivery. It was found that 11% of the mothers of fetuses that died and 2.5% of the mothers of fetuses that survived had had sexual intercourse during the period. It was concluded that intercourse during the month preceding delivery caused the fetal deaths. This conclusion:
- **a.** May be incorrect because mothers who had intercourse during the month before childbirth may differ in other important characteristics from those who did not
- **b.** May be incorrect because there is no comparison group
- **c.** May be incorrect because prevalence rates are used where incidence rates are needed
- **d.** May be incorrect because of failure to achieve a high level of statistical signi! cance
- **e.** Both *b* and *c*

3. All of the following are important criteria when making causal inferences *except:*
- **a.** Consistency with existing knowledge
- **b.** Dose-response relationship
- **c.** Consistency of association in several studies
- **d.** Strength of association
- **e.** Predictive value

Questions 4 and 5 are based on the following information.

Factor A, B, or C can each individually cause a certain disease without the other two factors, but only when followed by exposure to factor X. Exposure to factor X alone is not followed by the disease, but the disease never occurs in the absence of exposure to factor X.

4. Factor X is:
- **a.** A necessary and suf! cient cause
- **b.** A necessary, but not suf! cient, cause
- **c.** A suf! cient, but not necessary, cause
- **d.** Neither necessary nor suf! cient
- **e.** None of the above

5. Factor A is:
- **a.** A necessary and suf! cient cause
- **b.** A necessary, but not suf! cient, cause
- **c.** A suf! cient, but not necessary, cause
- **d.** Neither necessary nor suf! cient
- **e.** None of the above